

## METHOD FOR FABRICATING A NON-VOLATILE MEMORY

### CROSS-REFERENCE TO RELATED APPLICATION

This application claims the priority benefit of Taiwan application serial no. 91100555, filed January 16, 2002.

### BACKGROUND OF THE INVENTION

#### Field of Invention

5        [0001] The present invention relates to a method for fabricating a semiconductor device. More particularly, the present invention relates to a method for fabricating a non-volatile memory (NVM).

#### Description of Related Art

10       [0002] The non-volatile memory is widely used in many fields since the data stored in a non-volatile memory is retained when power is being not supplied to the non-volatile memory. The family of the non-volatile memory includes the mask read-only memory (Mask ROM) and the nitride read-only memory (NROM), which two have similar structures.

15       [0003] Refer to FIG. 1A~1D, FIG. 1A~1D illustrate a process flow of fabricating a non-volatile memory in the prior art in a cross-sectional view.

      [0004] Refer to FIG. 1A, a substrate 100 is provided. A strip stacked structure 101 comprising a gate oxide layer 102 (or a charge trapping layer 102), a polysilicon layer 104 and a nitride cap layer 105 is formed on the substrate 100. A buried drain

106 is then formed in the substrate 100 beside the strip stacked structure 101.

[0005] Refer to FIG. 1B, an insulating layer 108 is formed over the substrate 100 covering the strip stacked structure 101.

[0006] Refer to FIG. 1C, the insulating layer 108 is etched back to expose the nitride cap layer 105 and then the nitride cap layer 105 is removed. The insulating layer 108 thus covers only the buried drain 106.

[0007] Refer to FIG. 1D, a conductive layer 110 is then formed over the substrate 100 covering the polysilicon layer 104 and the insulating layer 108. The conductive layer 110 and the polysilicon layer 104 are then patterned successively to form a plurality of word-lines perpendicular to the buried drain 106 and a plurality of gates, respectively. The gates arranged in one row electrically connect with one word-line.

[0008] In the process described above, the conductive layer 110 usually comprises tungsten silicide ( $WSi_x$ ). However, since the tungsten silicon process requires a temperature higher than  $1000^{\circ}C$ , the dopants in the buried drain 106 easily diffuse out to shorten the channel between two buried drains 106. Therefore, the scalability of the memory device is restricted in consideration of the short channel effect (SCE). Moreover, the tungsten silicide word-lines have a higher resistance so that the performance of the memory device is hard to improve.

## SUMMARY OF THE INVENTION

[0009] Accordingly, this invention provides a method for fabricating a non-volatile memory to facilitate the scaling down of a memory device.

[0010] This invention provides a method for fabricating a non-volatile memory

to enhance the performance of the memory device.

10055265-012202

[0011] This invention provides a method for fabricating a NROM. A substrate having a strip stacked structure thereon is provided. The strip stacked structure comprises a conductive layer and a charge trapping layer, wherein the charge trapping layer can be a silicon oxide/silicon nitride/silicon oxide (ONO) stacked layer, a nitride/nitride/nitride (NNN) stacked layer, or a nitride/nitride/oxide (NNO) stacked layer. A buried drain is then formed in the substrate beside the strip stacked structure and an insulating layer is formed on the buried drain. A polysilicon layer and a cap layer are sequentially formed over the substrate. The cap layer, the polysilicon layer and the strip stacked structure are successively patterned in a direction perpendicular to the buried drain, wherein the strip stacked structure is patterned into a plurality of gates. A liner oxide layer is formed on the exposed surfaces of the gates, the substrate, and the polysilicon layer by thermal oxidation. Thereafter, the cap layer is removed, a metal layer is formed over the substrate, and then an annealing process is conducted to cause the metal to react with the polysilicon layer to form a metal silicide (self-aligned silicide) layer. The unreacted metal layer is then removed to leave the metal silicide layer to serve as a word-line that is electrically connected with the gates.

[0012] In the method for fabricating a NROM of this invention, the metal layer may comprise titanium (Ti) or cobalt (Co). When a titanium layer is adopted, the annealing process requires a temperature from about 600°C to about 800°C. When a cobalt layer is adopted, on the other hand, the annealing process requires a temperature from about 600°C to about 700°C.

[0013] This invention also provides a method for fabricating a Mask ROM, which is similar to the above-mentioned method for fabricating a NROM of this

invention, except that a gate dielectric layer, instead of the charge trapping layer, is formed on the substrate.

[0014] Since the temperature required in the metal salicide process in this invention is lower than that required in the tungsten silicide process in the prior art, the thermal budget is decreased and scaling down the memory device is therefore easier.

[0015] Moreover, since the resistance of the metal salicide (titanium/cobalt salicide) word-line in this invention is lower than that of the tungsten silicide word-line in the prior art, the performance of the memory device can be enhanced.

[0016] It is to be understood that both the foregoing general description and the following detailed description are exemplary, and are intended to provide further explanation of the invention as claimed.

#### BRIEF DESCRIPTION OF THE DRAWINGS

[0017] The accompanying drawings are included to provide a further understanding of the invention, and are incorporated in and constitute a part of this specification. The drawings illustrate embodiments of the invention and, together with the description, serve to explain the principles of the invention. In the drawings,

[0018] FIG. 1A~1D illustrate a process flow of fabricating a non-volatile memory in the prior art in a cross-sectional view; and

[0019] FIG. 2A~2L illustrate the process flow of fabricating a non-volatile memory according to a preferred embodiment of this invention.

#### DESCRIPTION OF THE PREFERRED EMBODIMENTS

[0020] Refer to FIG. 2A, a substrate 200 is provided and then a dielectric layer

202, a gate conductive layer 204, a nitride layer 206, and a patterned photoresist layer 208 are sequentially formed on the substrate 200. The gate conductive layer 204 comprises, for example, polysilicon. The dielectric layer 202 is a gate oxide layer or a charge trapping layer dependent on whether a Mask ROM or a NROM is being  
5 fabricated. When the dielectric layer 202 is a charge trapping layer, it comprises, for example, a silicon oxide/silicon nitride/silicon oxide (ONO) stacked layer, a nitride/nitride/nitride (NNN) stacked layer or a nitride/nitride/oxide (NNO) stacked layer.

[0021] Refer to FIG. 2B and FIG. 2C, wherein FIG. 2B schematically depicts a  
10 top view of a resulting structure formed after the subsequent manufacturing steps and FIG. 2C depicts a cross-sectional view of the resulting structure along the line I-I' in FIG. 2B. As that shown in FIG. 2B~2C, the nitride layer 206, the gate conductive layer 204, and the dielectric layer 202 are patterned with the photoresist layer 208 as a mask to form a plurality of strip stacked structures 207. As that shown in FIG. 2C, a  
15 liner layer 209 is formed on the exposed surfaces of the substrate 200 and the gate conductive layer 204. The liner layer 209 comprises, for example, silicon oxide and is formed by a method such as thermal oxidation.

[0022] Refer to FIG. 2D, a drain implantation 210 and a pocket implantation 212 are conducted to form a buried drain 214 and a pocket doped region (not shown),  
20 respectively, in the substrate 200 beside the strip stacked structure 207. The buried drain 214 is of N-type, for example.

[0023] Since the liner layer 209 is formed on the exposed surfaces of the substrate 200 and the gate conductive layer 204, damages will not be easily on the substrate 200 and the gate conductive layer 204 during the drain implantation 210 and

the pocket implantation 212.

[0024] Refer to FIG. 2E, an insulating layer 216 is formed over the substrate 200 covering the strip stacked structure 207 and the liner layer 209. The insulating layer 216 is formed by, for example, chemical vapor deposition (CVD) and comprises, for example, silicon oxide formed with tetraethyl-ortho-silicate (TEOS-oxide).

[0025] Refer to FIG. 2F, the insulating layer 216 is planarized with, for example, an etching-back process or a CMP process until the nitride layer 206 is exposed to leave the insulating layer 216 on the buried drain 214. The nitride layer 206 is then removed.

[0026] Refer to FIG. 2G, a silicon layer 218 and a cap layer 220 are sequentially formed over the substrate 200 covering the insulating layer 216 and the gate conductive layer 204. The silicon layer 218 comprises, for example, polysilicon and the cap layer 220 comprises, for example, silicon nitride.

[0027] Refer to FIG. 2H and FIG. 2I, wherein FIG. 2H schematically depicts a top view of a resulting structure formed after the subsequent manufacturing steps and FIG. 2I depicts a cross-sectional view of the resulting structure along the line II-II' in FIG. 2H. As that shown in FIG. 2H and FIG. 2I, a patterned photoresist layer 222, which has a plurality of strip patterns perpendicular to the buried drain 214, is formed on the cap layer 220

[0028] Refer to FIG. 2J, which depicts a cross-sectional view of the resulting structure formed after the subsequent manufacturing steps along the same line II-II' in FIG. 2H. The cap layer 220, the silicon layer 218, and the gate conductive layer 204 are patterned successively to form a plurality of gate structures 204a with the photoresist layer 222 as a mask. The photoresist layer 222 is removed. A liner layer

224 is then formed on the exposed surfaces of the silicon layer 218, the gate structure 204a and the substrate 200. The liner layer 224 comprises, for example, a silicon oxide layer formed by thermal oxidation.

[0029] Refer to FIG. 2K, the cap layer 220 is removed and then a metal layer 226 is formed over the substrate 200 covering the liner layer 224 and the silicon layer 218. The metal layer 226 comprises, for example, titanium (Ti) or cobalt (Co).

[0030] Refer to FIG. 2L, an annealing process is conducted to cause the metal layer 226 to react with the silicon layer 218 to form a metal silicide layer 226a. The unreacted metal layer 226 is then removed, leaving the metal silicide layer 226a to serve as a word-line electrically connecting with a plurality of gate structures 204a. The metal layer 226 may comprise titanium or cobalt. When a titanium layer is adopted, the annealing process requires a temperature from about 600°C to about 800°C. When a cobalt layer is adopted, on the other hand, the annealing process requires a temperature from about 600°C to about 700°C.

[0031] Since the liner layer 224 is formed on the exposed surfaces of the gate structures 204a and the substrate 200, the metal silicide layer 226a is not formed on the gate structures 204a nor on the substrate 200. Instead, the metal silicide layer 226a is formed on the silicon layer 218 located on the gate structures 204a.

[0032] Since the temperature required in the titanium/cobalt silicide process in the preferred embodiment of this invention is lower than the temperature required in the tungsten silicide process in the prior art, the thermal budget is decreased and the scaling down of the memory device is therefore easier. Moreover, since the resistance of the titanium/cobalt silicide word-lines in this invention is lower than that of the tungsten silicide word-lines in the prior art, the performance of the memory device can be

enhanced.

[0033] It will be apparent to those skilled in the art that various modifications and variations can be made to the structure of the present invention without departing from the scope or spirit of the invention. In view of the foregoing, it is intended that  
5 the present invention covers modifications and variations of this invention provided they fall within the scope of the following claims and their equivalents.

202210.595507  
1055265.012202